DOCUMENT RESUME

ED 091 743 CS 201 348

AUTHOR Gallo, Donald R.
TITLE Poetry Methods Rating Scale.
PUB DATE 68
NOTE 21p.; Reprinted from "Research in the Teaching of English," Fall 1968; See related documents CS 201 320-375

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS *Educational Research; Language Arts; *Measurement Instruments; *Poetry; Research Tools; Resource Materials; Secondary Education; *Teacher Attitudes; Teaching Methods; *Teaching Skills
IDENTIFIERS *The Research Instruments Project; TRIP

ABSTRACT

Designed to assess high school teachers' attitudes about teaching poetry, this questionnaire asked teachers to respond to a 38-item poetry methods rating scale (PMRS) on a seven-point scale (from "strongly agree" to "strongly disagree"). The items for the questionnaire were derived from a study of popular methods texts for teaching literature. Scores on the questionnaire were compared with the scores originally obtained from experts in teaching English. The Spearman-Brown split-halves reliability for 39 teachers in the study in which the PMRS was used was .75. The test-retest reliability coefficient was .62 for 93 other teachers on whom the questionnaire was tried out. [This document is one of those reviewed in The Research Instruments Project (TRIP) monograph "Measures for Research and Evaluation in the English Language Arts" to be published by the Committee on Research of the National Council of Teachers of English in cooperation with the ERIC Clearinghouse on Reading and Communication Skills. A TRIP review which precedes the document lists its category (Teacher Competency), title, author, date, and age range (high school), describes the instrument's purpose and physical characteristics.] (JM)

ED 091743

NCTE Committee on Research

The Research Instruments Project (TRIP)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The attached document contains one of the measures reviewed in the TRIP committee monograph titled:

Measures for Research and Evaluation in the English Language Arts

TRIP is an acronym which signifies an effort to abstract and make readily available measures for research and evaluation in the English language arts. These measures relate to language development, listening, literature, reading, standard English as a second language or dialect, teacher competencies, or writing. In order to make these instruments more readily available, the ERIC Clearinghouse on Reading and Communication Skills has supported the TRIP committee sponsored by the Committee on Research of the National Council of Teachers of English and has processed the material into the ERIC system. The ERIC Clearinghouse accession numbers that encompass most of these documents are CS201320-CS201375.

TRIP Committee:

W.T. Fagan, Chairman
University of Alberta, Edmonton

Charles R. Cooper
State University of New York at Buffalo

Julie M. Jensen
The University of Texas at Austin

Bernard O'Donnell
Director, ERIC/RCS

Roy C. O'Donnell
The University of Georgia
Liaison to NCTE Committee on Research

CS 201 348

NATIONAL COUNCIL OF TEACHERS OF ENGLISH
1111 KENYON ROAD
URBANA, ILLINOIS 61801

Category: Teacher Competency

Title: Poetry Methods Rating Scale

Author: Donald R. Gallo

Description of the Instrument:

Purpose: To assess high school teachers' attitudes about teaching poetry.

Date of Construction: 1968

Physical Description: With the thirty-eight item PMRS teachers are asked to respond on a seven point scale (from strongly agree, to strongly disagree) to statements like the following: "Before the class reads and studies a poem, the teacher should tell the students to look or listen for specific things." The items for the questionnaire were derived from a study of popular methods texts for teaching literature. The original pool of items was first tried out on "experts in teaching English," and their scores on each of the retained items in the final form of the questionnaire permit comparison of the teachers' scores with the experts' scores.

The questionnaire can be completed in about twenty minutes.

Validity, Reliability, and Normative Data:

The Spearman-Brown split-halves reliability for thirty-nine teachers in the study in which the PMRS was used was .75. The test-retest reliability coefficient was .62 for ninety-three other teachers on whom the questionnaire was tried out. The author concludes: "The PMRS is a reliable instrument for assessing English teachers' opinions of methods of teaching poetry."

Evidence for validity was sought through a number of correlations. Positive and statistically significant ones were found between PMRS scores and the Teaching Situation Reaction Test, students' evaluations of their

teachers, years of teaching experience, and amount of poetry read and enjoyed by students in a teacher's classes. In addition, the experts' screening of the items contributes to validity. The author concludes that while evidence of validity is not strong, some item revision and tryouts on a larger sample will probably produce higher validity coefficients.

Experts' scores on the questionnaire and their scale scores on individual items provide a kind of normative data to which teachers' scores can be compared. In addition, the reported mean scores (by years of teaching experience) for the upstate New York teachers in the study provide further normative data.

Ordering Information:

EDRS

Related documents:

Also available in Donald R. Gallo, "Toward a More Effective Assessment of Poetry Teaching Methods," Research in the Teaching of English, 2 (Fall 1968), 125-141.

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
National Council of Teachers of English
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

RESEARCH IN THE TEACHING OF ENGLISH

VOLUME 2, NUMBER 2, FALL 1968

Contents

ARTICLES

93 A study of the Hawthorne effect in educational research
by Desmond L. Cook and JoAnn King

99 Teacher questioning behavior in nine junior high school English classes
by James Hoetker

107 Comment on the Hoetker report
by James R. Squire

109 Beyond the ETPS guidelines: methods course content
by Donald D. Small

115 The proper study
by Nancy S. Boze

125 Toward a more effective assessment of poetry teaching methods
by Donald R. Gallo

142 The effect of reading for a creative purpose on student attitudes toward a short story
by Richard J. Smith

152 A comparison of words in color and traditional basal reading
by William G. Dodds, Jr.

161 Transformational grammar and the teaching of reading
by Joseph C. Beaver

Mr. Gallo's research specifically focuses on the evaluation of poetry teaching methods, but the scale he is developing may well offer a model for the evaluation of methods for teaching other aspects of English, too.

Toward a more effective assessment of poetry teaching methods[1]

DONALD R. GALLO
University of Colorado, Denver Center

Hundreds of studies and dozens of rating scales to measure teaching effectiveness have been produced in the twentieth century. Yet most measuring devices are extremely subjective and unreliable. This is not to say that important advances have not been made. Some of the most recent instruments and procedures, such as those developed by Hughes,[2] Medley and

---

[1] This paper is based on a doctoral dissertation accepted by the Graduate School of Syracuse University, June, 1968, under the title "The construction and validation of an instrument to assess teachers' opinions of methods of teaching English to tenth grade students of average ability." The complete dissertation is available from University Microfilms. The author is indebted to Margaret J. Early, John B. Hough, William W. West, and the numerous English educators and teachers who contributed to the study.

[2] Marie Hughes and others, *Development of the means for the assessment of the quality of teaching in elementary schools* (U.S. Office of Education Cooperative Research Project No. 353. Univer. of Utah, 1959).

Mitzel,[3] Ryans,[4] Bellack,[5] and Flanders,[6] and modifications of Flanders' Scale,[7] have been more successful than previous attempts. Nevertheless, almost all of the major studies in the field have been concerned with general teaching procedures, such as maintaining order, giving information, rewarding answers and manipulating discussion, and with factors related to teaching, such as voice, appearance, personal interests, and academic background. Moreover, these factors have been applied in the same way to teachers in diverse academic fields. Few studies have attempted to evaluate teachers in one specific field with criteria from that field.

The most recent efforts to develop instruments for assessing teaching effectiveness have been concerned with coding and analyzing teaching behavior as it occurs in the classroom. But here, too, no distinction has been made in the teaching style of English teachers as compared with social studies or science teachers. Neither has a distinction been made between the styles used for teaching poetry and those used for teaching short fiction or writing or grammar. But analyzing teaching behavior is only one way of evaluating teachers.

Another evaluation procedure is to assess teachers' knowledge and opinions of various academic subjects. But this pursuit fell into neglect years ago because the early studies—of Meriam[8] and Jones[9] for example—indicated no significant relationships between teachers' knowledge of academic subjects

---

[3] D. M. Medley and H. E. Mitzel, "A technique for measuring classroom behavior," *Journal of Educational Psychology*, 1958, 49, 86-92.

[4] D. G. Ryans, *Characteristics of teachers* (Washington, D.C.: American Council on Education, 1960).

[5] A. A. Bellack, J. R. Davitz, and others, *The language of the classroom* (New York: Institute of Psychological Research, Teachers College, Columbia University, 1963).

[6] N. A. Flanders, *Teacher influence, pupil attitudes, and achievement studies in interaction analysis* (Office of Education Cooperative Research Project No. 379. Washington, D.C.: U.S. Department of Health, Education, and Welfare, 1960).

[7] See E. J. Amidon and J. B. Hough (eds.), *Interaction analysis: theory, research, and application* (Reading, Mass.: Addison-Wesley, 1967).

[8] J. L. Meriam, *Normal school education and efficiency in teaching* (Contributions to Education, No. 1. New York: Bureau of Publications, Teachers College, Columbia Univer., 1906).

[9] R. D. Jones, "The prediction of teaching efficiency from objective measures," *Journal of Experimental Education*, 1948, 15, 85-100.

and various measures of teaching success. However, it is quite possible that those measures of subject matter knowledge were too general to reveal any significant relationships with teaching efficiency. Moreover, knowledge of the *subject matter* of English is quite different from knowledge of *methods of teaching* that subject matter.

Although the secondary school English methods texts are filled with recommendations for teaching various aspects of the subject, no one has studied statistically the quality of the teacher who uses the recommended methods and the teacher who does not. Tovatt and others at Ball State University produced a list of 30 statements of methods of teaching English[10] that was backed by a "rationale" of quoted theory and research.[11] But their list is very broad and has been recommended as a self-rating device rather than as a validated instrument for evaluating English teaching.

PURPOSE

Based on the premise that knowledge of English teaching methods is a key element in teaching success, the author attempted to construct an instrument for assessing English teachers' knowledge of poetry teaching methods. But knowledge of methods is difficult to define and consequently to measure. Constructing a test to measure knowledge of English teaching methods is further complicated by the lack of conclusive research as to what methods are best for teaching what content at what stages of development. Therefore, the instrument was constructed to measure teachers' *opinions* of certain methods of one aspect of English—teaching poetry. In addition to limiting the scope of the instrument to methods of teaching poetry, it was limited still further for control purposes to teaching poetry to tenth grade students of average ability.

The opinions of the teachers under examination were assumed to represent "professional" opinions in the sense that they were based on facts and teaching experiences and were not merely unfounded, unverifiable opinions. It was also necessary to assume that the opinions teachers expressed on the instrument were their honest opinions, though they might

---

[10] A. L. Tovatt and others, "A sampler of practices in teaching junior and senior high school English" (Champaign, Ill.: NCTE, 1965).

[11] A. L. Tovatt and others, "Rationale for a sampler of practices in teaching junior and senior high school English" (Champaign, Ill.: NCTE, 1965).

possibly have been what the teachers thought the investigator expected them to say. But this is an unfortunate limitation of almost all written responses.

As a single instrument, the validity is dependent upon teachers' opinions of poetry teaching methods. But those opinions also operate in teaching behavior. In turn, that teaching behavior is relatable to other measures. It is relatable to non-substantive, school-oriented measures—identified by the Minnesota Teacher Attitude Inventory[12] and the Teaching Situation Reaction Test[13]—and to non-substantive, non-school-oriented measures, identified by the Rokeach Dogmatism Scale.[14] Construct validity of the poetry scale was investigated by means of these three instruments. Criterion validity was investigated through measures of teaching content: poetry lessons taught by teachers, a poetry test taken by students and an evaluation of the teachers made by the students. The reliability of the instrument was established by a test-retest procedure.

Thus, a study was designed to construct an instrument—called the Poetry Methods Rating Scale (PMRS)—for assessing teachers' opinions of methods of teaching poetry to tenth grade average ability students and to validate it by determining the relationship between scores on the instrument and teachers' attitudes, personality, performance, and success in the classroom.

Hypotheses

On the basis of what the Poetry Methods Rating Scale was expected to measure, the following hypotheses were formed

H_1 The PMRS will be positively correlated with the MTAI

H_2 The PMRS will be positively correlated with the TSRT

H_3 The PMRS will be positively correlated with the relative openness of belief-disbelief systems as measured by the Dogmatism Scale.

H_4 The PMRS will be positively correlated with teachers' scores for teaching three poems.

H_5 The PMRS will be positively correlated with students' scores on a test of three poems.

---

[12] W. W. Cook, C. H. Leeds, and R. Callis, *Minnesota Teacher Attitude Inventory* (New York: The Psychological Corporation, 1951).

[13] J. K. Duncan and J. B. Hough, "Technical review of the Teaching Situation Reaction Test" (Unpublished mimeographed manuscript, Ohio State University, September, 1966).

[14] M. Rokeach, *The open and closed mind* (New York: Basic Books, 1960).

H_6 The PMRS will be positively correlated with students' evaluations of their teachers.

H_7 The PMRS will be positively correlated with a retest on the same instrument when no poetry teaching has intervened.

H_8 The PMRS will be positively correlated with a retest on the same instrument after three poems have been taught.

PROCEDURES

Experimental Instrument

The author constructed a 7-category equal-appearing interval scale consisting of 62 statements about methods of teaching poetry. For example: "When a new poem is introduced, the teacher should first read it aloud to the class." Or "A good way to begin the study of poetry in tenth grade is by defining the word *poetry*." Items were written on the basis of recent research and theory reported in the journals and methods texts, especially the most recent texts written by Hook, Loban, Ryan and Squire, Fowler, Burton, Dunning, and Sweetkind.

The directions ask teachers to react to each statement on the instrument as it would apply to a tenth grade class of average ability students (i.e., neither the very bright nor the very dull). Using their opinions of what are good and poor methods of teaching poetry, teachers are to indicate how strongly they agree or disagree with each statement in terms of seven categories from *Strongly Agree* through *Neutral* to *Strongly Disagree*.

The 62-item instrument was sent to 45 experts in teaching English—including writers of methods texts, members of the NCTE Poetry Committee, and other selected English educators. Those experts were asked to respond to each item in terms of agreement or disagreement on the 7-point scale. From the responses of the 32 experts who returned the scale, Q-scores and scale-scores were computed for each item according to formulas described by Edwards in *Techniques of Attitude Scale Construction*,[15] which was based largely upon the 1929 work of Thurstone and Chave.[16]

Q—or interquartile range—is a measure of dispersion of the ratings of the middle 50% of the judges. In another sense, Q

---

[15] A. L. Edwards, *Techniques of attitude scale construction* (New York: Appleton-Century-Crofts, 1957), pp. 83-119, 149-171.

[16] L. L. Thurstone and E. J. Chave, *The measurement of attitude: a psychophysical method and some experiments with a scale for measuring attitude towards the church* (Chicago: The University of Chicago Press, 1929).

is a measure of the ambiguity of an item. The smaller the Q-score, the less ambiguous a statement is and the better the reliability is among the judges. Items on which the Q-score was 2.00 or greater were eliminated from the final form of the Poetry Methods Rating Scale, leaving 38 items.[17] The use of 2.00 as a cut-off point limited the dispersion of the middle 50% of the ratings and created a conservative but probably more reliable scale.

The scale-score of each item is the median score of the experts' ratings. Rounded to the nearest whole number, each scale-score becomes the "right" answer for the item. A teacher's score for an item on the final scale is computed by subtracting one point for each deviation on either side of the correct category. A teacher's opinion is therefore evaluated in terms of its deviation from the average opinion of a group of 32 experts. The Spearman-Brown split-halves (odd-even) reliability of the experts' scores was .772. The mean scale-value was 4.09; the mean Q-value was 1.54, with a mean Q of 1.55 for the odd numbers and 1.52 for the even numbers. These scores indicate a relatively good balance of both scale-scores and Q-scores. Half the mean Q-values of all judgments provides an indication of how reliable the scale-values are; that value was .77 for this scale.

Population

A total of 39 tenth grade English teachers of average ability students were tested from 14 schools in 11 districts in Central New York State. Of the 12 males and 27 females in the study, 24 had a bachelor's degree and 15 had a master's degree or its equivalent. Teaching experience ranged from no previous experience to 35 years, with a mean of 5.18 years. The distribution is shown in Table 1. The number of teachers from urban

Table 1

Teaching Experience

Years	N	Years	N
0	7	5	0
1	6	6-10	6
2	5	11-18	0
3	6	19-35	4
4	5		

[17] The final form of the PMRS is reproduced at the end of this article.

schools was 16, with 15 from suburban and 8 from rural schools.

Only one average ability tenth grade class from each teacher was involved in the study, making a total of 851 students. Average ability classes were defined as neither the very bright nor the very dull, having a mean IQ of between 100 and 110.

Measuring Instruments

Three instruments were used in the study: the Minnesota Teacher Attitude Inventory (MTAI), which measures a teacher's ability to maintain harmonious relationships in the classroom and is a reliable predictor of teaching success in the realm of human relations; the Teaching Situation Reaction Test (TSRT), a relatively new test which measures factors indirectly related to general teaching performance by evaluating teachers' reactions to various teaching situations which the test poses; and the Rokeach Dogmatism Scale, which measures the relative openness or closedness of a person's belief-disbelief systems. The investigator anticipated that the factors these tests measure might be related to the factors measured by the Poetry Methods Rating Scale.

Collecting Data

Each of the 39 teachers who had agreed to participate in the study was tested on the PMRS, MTAI, TSRT, and D-scale. The teachers also filled out an information sheet that asked for degrees held, years of teaching experience, major subject in college, and whether they had taken a methods course, done student teaching, etc.

On the basis of their scores on the Poetry Scale, the teachers were divided into two groups: those above the median and those below. Thirteen teachers from the high-scoring group were randomly selected along with 12 from the low-scoring group. Those 25 teachers were asked to teach three short poems to one of their tenth grade average ability classes at their convenience sometime during a four-week period designated by the investigator. They were instructed to teach the poems in any order and in any way they chose, providing they taught the poems with an aim to increasing the students' understanding and appreciation of poetry generally and of those poems in particular. The poems were "Sonic Boom" by John Updike, "A Coney Island Life" by James L. Weil—both contemporary poets—and "God's Grandeur" by Gerard Manley Hopkins. Only 21 teachers completed this part of the study.

Each teacher was asked to tape-record his lessons on each

of the three poems. The taped lessons were then evaluated by the investigator and two others experienced in the evaluation of English teachers. Each lesson was given a general evaluation by each evaluator. However, although this was a general evaluation, it was not based merely on a general impression or "feeling" about the lesson. Each lesson was examined in terms of the teacher's general organization, introduction, interpretation of meaning, discussion of form, examination of language, and use of related activities. Each lesson received a rating of from 1 (for poor) to 5 (for excellent) from each rater. Each teacher's total score was the sum of the scores of the individual lessons from each of the three evaluators, allowing for a total perfect score of 45. Inter-rater reliability coefficients ranged from .50 to .85 with an average of .72.

In order to assess the achievement of the students, each teacher administered a multiple-choice test after teaching the three poems. The test, constructed by the investigator, covered the most significant elements of meaning, form, language, and style of each of the three poems and was designed to measure the kinds of things an excellent teacher would be expected to teach. The teachers did not have access to the poetry test until they finished teaching the poems.

As a measure of the poetry test's ability to assess the effect of teaching the three poems, the same poems and test were given to one tenth grade average ability class of each of the 14 control teachers who did not teach the poems. Those students were instructed to read the poems and take the test. In both the experimental and control groups, the students could refer to the poems as they took the test.

As a final step in evaluating the effectiveness of the teachers, each student in the experimental group was asked—in a short questionnaire—anonymously to evaluate his teacher. The questionnaire consisted of eight questions about the quality of the teacher's poetry teaching as well as the student's own enjoyment and learning related to the three poems. This questionnaire was administered personally by the investigator at the end of the class period in which the students had taken the poetry test. The main question used to evaluate the teachers was "In comparison with all other English teachers, how would you rate your English teacher *as a teacher of poetry*? a. superior, b. average, c. below average." Each response was scored from 3 (for superior) to 1 (for below average). A

median score was computed for each class' evaluation, then multiplied by 100 to eliminate decimals.

After each teacher completed his teaching, he was asked to take the Poetry Methods Rating Scale a second time. The control teachers were retested as well. In addition, 93 secondary English teachers from most of the schools participating in the study were given the PMRS as a test and retest after an interval of three to four weeks.

RESULTS

The scores of the 39 teachers in the study ranged from 200 to 241 on the PMRS, with a mean of 221.795 and a standard deviation of 10.024. The Spearman-Brown split-halves reliability of their scores was .749. The scores of the 93 teachers used to establish reliability ranged from 190 to 242 with a mean of 222.118 and a standard deviation of 9.342.

An informal item analysis indicated that 8 items discriminated poorly, 16 discriminated adequately, and 14 discriminated extremely well.

Because the study was concerned with a new instrument, a relatively small sample, and a procedure that involved unaccountable variables of teacher personality and experience, the investigator chose .10 as an acceptable level of significance with a two-tailed test of hypotheses. This choice, instead of the conventional .05 or .01 levels, would then indicate directionality where the data were not convincingly significant at a higher level.

In testing the construct validity of the PMRS, no significant correlation was found between the PMRS and the MTAI (.067) nor between the PMRS and the Dogmatism Scale (.092). A significant (.05 level) Pearson product-moment correlation coefficient of .367 resulted between PMRS and TSRT scores. The PMRS therefore does not measure teaching attitudes or open-mindedness, but it does measure some factors indirectly related to general teaching performance.

Tests of the three hypotheses designed to establish criterion validity were rejected. However, a *t*-test of the difference between means resulted in a significant difference (.10 level) in students' evaluations of teachers, favoring teachers scoring above the mean on the PMRS. Although there were no significant differences in teachers' scores for teaching or in students' poetry test scores, the teachers scoring above the mean on the Poetry Scale had slightly higher mean scores. (See Table 2)

Table 2
Criterion Validity
Mean Scores

Variable	Above M.	Below M.	r
Teaching Scores	22.33	20.22	.301
Poetry Test	20.44	19.95	.136
Pupil Evaluation	245.92	225.50	.411

A multiple regression correlation coefficient of .667 resulted from a combination of the PMRS, TSRT, MTAI, teaching score, and students' evaluation. An analysis of variance on that data resulted in a F of 3.211, significant at the .05 level.

No significant differences were found in PMRS scores according to sex, degrees held, type of school, or various educational variables: whether or not teachers had been English majors in college, had taken a course in English methods, or had done student teaching. However, teachers with a master's degree were twice as likely to score above the mean, and rural teachers were more likely to be below the mean on the PMRS.

The most interesting differences in mean scores on the PMRS occurred on the basis of amount of teaching experience. Table 3 indicates that the highest mean score was obtained by teachers with from six to ten years of experience. A close second was teachers with no previous experience. The group scoring lowest had from 19 to 35 years of experience. The second lowest-scoring group had one year of experience. Teachers with from two to four years of experience were at or

Table 3
Teaching Experience
Mean Scores

No. of years*	N	M.
0	7	226.286
1	6	217.833
2	5	221.600
3	6	220.000
4	5	222.400
6-10	6	227.333
19-35	4	213.750

*Not counting the present year.

near the mean. The two high-scoring groups differ significantly from the two low-scoring groups at the .10 level.

The trend favoring high-scoring teachers is seen in the reactions of students on the evaluation questionnaire. Students in classes of teachers who scored above the mean on the PMRS did not differ significantly from students in classes of teachers below the mean, not in their liking for poetry in general, or in their preference for any one of the three poems, or in their belief that they learned more about poetry from studying those poems. However, the students of the high-scoring teachers had read more poetry in class during the year, enjoyed studying the three poems more, and rated their teachers significantly higher than students of teachers scoring below the mean on the PMRS.

The test-retest reliability coefficient was .624 for teachers who taught the poems, .697 for the control teachers who did not teach, and .624 for the 93 other teachers.

DISCUSSION

One possible reason for the low correlation between the PMRS and teachers' scores for teaching the poems is the great variance among the teachers, especially the high-scoring teachers. Moreover, the three evaluators concluded that most of the poetry teaching was not very effective, even from the teachers scoring high on the other variables in the study. On the basis of the criteria of excellent teaching agreed upon by the three evaluators, much of the teaching was below average. In fact, the average score was only 21.4 out of a possible 45 points, or only 2.4 points per poem. On the 5-point scale used, no evaluator gave a 5, and there were few 4's. Most of the teachers lectured most of the time, elicited few student comments, progressed line-by-line through the poems without starting with general impressions and then discussing the elements of the poems which led to those impressions, and made little effort to teach the skills of poetry reading or interpretation.[18]

In light of the relatively poor teaching, it is no surprise that the correlation between teaching scores and students' poetry test scores was only .157. The poetry test had been designed to evaluate elements of poetry that an excellent teacher would have focused on, but no teacher met the three evaluators' criteria for excellent teaching. Students who scored high

[18] Editor's note: Perhaps the fact that their lessons were being tape recorded led some teachers to adopt more routine methods than they usually employ.

on the test seemed to do so on their own ability, because they had little stimulation or direction from their teachers.

What is surprising is that the correlation between teachers' scores on the PMRS and their scores for teaching the poems was only .301. There is clearly a trend here as well as in much of the data in favor of the teachers scoring high on the PMRS, but it is quite obvious that what teachers know and believe—or at least say they know and believe—about methods of teaching poetry does not always result in related behaviors in their classes. For example, item #7 on the PMRS states "Students should be urged to defend their interpretations of poems by quoting passages from the poems." The experts strongly agree with that statement. Eight of the 21 teachers who taught the poems agree with it and 9 strongly agree. Thus, 17 of the 21 teachers agree that students should defend their interpretations of poems; yet the taped lessons revealed very few instances of students supporting their interpretations. Therefore, although this study presents evidence to support the contention that teachers' knowledge and beliefs about methods of teaching poetry have a direct bearing on how they teach, there is not a one-to-one relationship by any means. In some instances, as in the example noted above, there seems to be almost no relationship whatever, although every one of the four teachers who disagreed with statement #7 scored below the mean on the PMRS. Nevertheless, when individual items are combined into a rather comprehensive instrument as they are in the Poetry Scale, there is evidence of a relationship between teachers' opinions and other measures of their efficiency in the classroom.

In spite of the relatively low correlations between the PMRS and the independent variables used in this study, when certain of those instruments and procedures are combined they yield a relatively high multiple correlation. The highest correlation resulted from a combination of the PMRS, TSRT, MTAL teaching score, and students' evaluation. This indicates that the combination of those five variables produces a more valid profile of English teachers than any one of them does individually. Moreover, it suggests that, when coupled with certain of the independent variables used in this study, the PMRS is a more powerful instrument than it appears to be alone.

The primary implication of this research is that it is possible to construct a valid and reliable instrument to measure

CONCLUSI

teachers' opinions of methods of teaching poetry, and probably to expand this kind of scale to measure opinions of methods of teaching other aspects of English. This scale clearly measures different aspects of teaching effectiveness than do the TSRT and the MTAI. It is also suspected that it measures aspects of teaching that differ considerably from those aspects measured by interaction analysis-type scales, though a separate study will have to be carried out to test this belief.

Associated very closely with the development of the PMRS were the ratings made by the groups of experts in English education. As previously noted, it is difficult to measure teachers' knowledge of methods because there is a lack of specific research and agreed-upon theory on which to base such a scale. The development of the PMRS has provided a list of concrete statements about methods of teaching poetry with a corresponding list of expert opinions about those methods. This study has shown, for example, that most experts strongly disagree with the statement: "The study of metrics should be one of the first steps in approaching poetry in tenth grade," although two of the 32 experts surveyed agree with that. At the same time, by examining those items with a Q-value in excess of 1.99 on the preliminary form of the PMRS, it is possible to see which methods experts did not agree on. For example, the statement "Groups of students should be allowed to discuss different elements of the poem in differentiated groups in the same classroom at the same time" brought a wide variety of responses from the experts. One would expect from the number and tone of recent articles on the values of grouping, that most of the experts would have agreed to that statement. Instead, 18 agreed, 6 were neutral, and 7 disagreed. The same was true of experts' opinions regarding the value of pre-teaching vocabulary: "Before introducing a new poem, the teacher should teach a few of the important vocabulary words the students will encounter." In any case, the PMRS has provided a compilation of experts' opinions about specific methods of teaching poetry.

CONCLUSIONS

The results indicate that—in light of the limitations and conditions set forth in this pilot study—the Poetry Methods Rating Scale is a reliable instrument for assessing English teachers' opinions of methods of teaching poetry. The validity, however, is tenuous. The ratings of the experts on the PMRS are certainly valid. There is an acceptable correlation between the

PMRS and the TSRT. There are significant relationships between the PMRS and students' evaluations of their teachers, years of teaching experience of teachers, and amount of poetry read and enjoyed by students. However, a number of other key variables—especially poetry teaching scores and poetry test scores—were not significantly correlated with the PMRS. Nevertheless, the data show consistent trends in many cases to support the contention that, with certain modifications in the scale as well as the research procedures, significant correlations may result in future validity studies.

There is, then, a need for much improvement before the Poetry Methods Rating Scale can be used with confidence to assess English teachers' opinions of poetry teaching methods. The item analysis indicates the need to eliminate or revise 8 of the 38 items. In addition, there is reason to suspect that a larger sample will lead to more significant t values since so many of the differences between means of teachers scoring above the mean and those scoring below the mean on the PMRS were not significant but favor the high-scoring teachers. The revision of selected items and the use of a larger sample may well produce higher validity coefficients.

Poetry Methods Rating Scale

On the following pages are some statements about methods for teaching poetry. React to each statement as it would apply to a tenth grade class of average ability students (i.e. they are neither the very bright nor the very dull). Using your opinions of what are good and poor methods for teaching poetry, mark on the separate answer sheet how strongly you agree or disagree with each statement in terms of the following seven categories:

Strongly agree
Agree
Agree, with some exceptions
Neutral—sometimes agree, sometimes disagree
Disagree, with some exceptions
Disagree
Strongly disagree

For example, how strongly do you agree or disagree with the following method?

"A good way to begin the study of poetry in tenth grade is by defining the word *poetry*."

If you agree to that statement without exception, you would make a check on the answer sheet in the column marked *Strongly*

Agree. If you disagree with it but might agree in some cases, check the column marked *Disagree, with some exceptions.*

Proceed through the items in order, marking only *one* choice for each statement. Do not omit any items.

[Editor's note: The median score of the experts' rating is given in brackets after each item below, though, of course, that did not appear on the form of the PMRS administered to the teacher.]

1. The teacher should lead the students from the simple to the complex in a poem—starting with the who, what, when, where and progressing to the symbolic. [5.33—Agree,with exceptions]
2. The teacher should ask the students to identify the form and mechanics (meter, rhyme, figures of speech, etc.) in each poem. [1.93—Disagree]
3. Poem studied in tenth grade should be chosen for their appeal to the senses and emotions of the students. [4.84—Agree, with exceptions]
4. Poetry in tenth grade should be studied as a unit by itself. [3.87—Neutral]
5. The teacher should use recordings of poems to help tenth grade students appreciate the sounds of poems. [6.11—Agree]
6. The main interpretation of a poem should be based on the poem itself. [6.70—Strongly agree]
7. Students should be urged to defend their interpretations of poems by quoting passages from the poems. [6.50—Strongly agree]
8. Tenth grade students should first understand the literal meaning before discussing the symbolic meaning of a poem. [6.24—Agree]
9. Each teacher should decide which poems will be read and studied in his class. [4.86—Agree, with exceptions]
10. Tenth grade students should be asked to define and identify various verse forms: quatrains, heroic couplets, blank verse, Shakespearean and Petrarchan sonnets, etc. [2.60—Disagree, with exceptions]
11. The teacher should require the students to write a prose paraphrase (although not from memory) of each poem studied in class. [1.79—Disagree]
12. Before the class reads and studies a poem, the teacher should tell the students to look or listen for specific things. [4.23—Neutral]
13. The teacher of poetry should read widely in the fields of poetry and literary criticism. [6.74—Strongly agree]
14. The study of every poem should culminate in a statement of its message. [1.66—Disagree]

15. After a poem has been thoroughly discussed in class, the teacher should summarize the main points. [3.99—Neutral]
16. Poetry should be studied primarily because of its importance as a literary genre. [1.71—Disagree]
17. When poetry is studied in tenth grade, students should be assigned about four to six new poems to read for homework each night. [1.65—Disagree]
18. The poems used in class should appeal to the immediate needs and interests of the students. [4.71—Agree, with exceptions]
19. Each student should be required to recite a poem in front of class. [1.29—Strongly disagree]
20. Topics for writing during a poetry unit should be related to the subjects of the poems being read and discussed in class. [4.47—Neutral]
21. The mechanics of poetry should be studied to see where and how they contribute to the meaning of particular poems. [5.69—Agree]
22. Students should give the one correct interpretation of each poem in order to receive full credit for their answers on tests. [1.26—Strongly disagree]
23. Important facts of a poet's life and times should be introduced only when they have some relevance to a particular poem being studied. [6.06—Agree]
24. A good way to begin the study of poetry in tenth grade is by reading a few short, humorous poems. [4.62—Agree, with exceptions]
25. It is better to examine only a couple of poems in close detail than to examine a greater number of poems adequately. [3.76—Neutral]
26. When poetry is studied in tenth grade, the mechanics (meter, rhyme, figures of speech, etc.) should receive as much or even more attention than the meaning of individual poems. [1.50—Disagree]
27. Pleasure should precede analysis of poems. [5.96—Agree]
28. Students should be given the opportunity to participate in choral readings. [5.84—Agree]
29. As part of the study of language in tenth grade, such things as word meanings, denotations and connotations, word histories, and word order should be examined in poems. [4.96—Agree, with exceptions]
30. There is little time for modern folk songs and ballads in the tenth grade curriculum. [1.66—Disagree]
31. In addition to other work with sentence patterns, students should study these patterns as part of a poetry unit by examining the word order in poems. [5.05—Agree, with exceptions]

32. Students should be given the freedom to read only those poems or types of poems they want to read. [2.38—Disagree]
33. The study of metrics should be one of the first steps in approaching poetry in tenth grade. [1.22—Strongly disagree]
34. Students should be asked to try their hand at *haiku* as an early step in writing poetry in tenth grade. [4.54—Agree, with exceptions]
35. Poetry should be studied primarily for its vivid recreation of human experience. [6.36—Agree]
36. One of the main goals of poetry study should be for the students to learn the facts about the life and times of the poets, such as important dates and main events. [1.35—Strongly disagree]
37. Students in tenth grade should be asked to read a favorite poem to the class, but only after individual preparation for oral reading. [5.14—Agree, with exceptions]
38. With complicated poems, more than one interpretation should be allowed. [6.40—Agree]